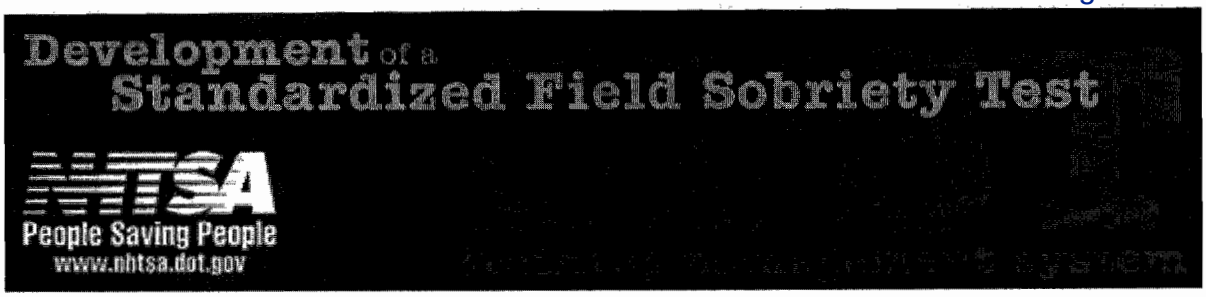

Introduction

This report presents the results of a study conducted for the National Highway Traffic Safety Administration to develop a model system for managing Standardized Field Sobriety Test (SFST) training within a state. The report is presented in three sections. This brief Introduction discusses the historical context of the study and presents the objectives of the research. The second section of the report describes the tasks that were performed and presents study results. The third section discusses the specifications of a computerized model system for tracking SFST training.

Development and Validation of the SFST Battery

During the late 1960s and early 1970s more than 50,000 people lost their lives each year on our nation's public roads; more than half of the fatalities involved an alcohol-impaired driver. Traffic safety has improved considerably since that time: the annual death toll has declined to about 40,000, even though the numbers of drivers, vehicles, and miles driven all have greatly increased. When miles traveled are considered, the likelihood of being killed in traffic in 1966 was more than three times what it is today.

Research sponsored by the National Highway Traffic Safety Administration (NHTSA) has contributed to the improved condition, in part, by providing law enforcement officers with useful and scientifically valid information and training materials to assist in the enforcement of drinking and driving laws. Beginning in 1975, NHTSA sponsored research that led to the development of a Driving While Impaired (DWI)[1] detection guide that listed 20 driving cues and the probabilities that a driver exhibiting a cue would have a BAC of at least 0.10 percent (Harris et al., 1980; Harris, 1980). A similar study was conducted more recently that identified 24 driving cues that are predictive of DWI at the 0.08 level (Stuster, 1997); the latter study also identified ten post-stop cues with probabilities of DWI of at least 90 percent. NHTSA previously sponsored research that led to the development of a motorcycle DWI detection guide and training program (Stuster, 1993). NHTSA's DWI training materials, based on the results of these studies, have exposed the current generation of law enforcement officers in the U.S. to information critical to DWI enforcement by providing a systematic, scientifically valid, and defensible approach to on-the-road DWI detection.

At the same time NHTSA was providing officers with information concerning the driving behaviors that are the most predictive of impairment, the agency also sponsored research that led to the development of a standardized battery of tests for officers to administer to assess driver impairment after an enforcement stop has been made. Marcelline Burns and Herbert Moskowitz conducted laboratory evaluations of several of the tests that were most frequently-used by law enforcement officers at the time (Burns and Moskowitz, 1977). In addition to a variety of customary roadside tests (e.g., finger-to-nose, maze tracing, backward counting), the researchers evaluated measures of an autonomic reaction to central nervous system depressants, known as Horizontal Gaze Nystagmus. Horizontal Gaze Nystagmus (HGN) is an involuntary jerking of the eye that occurs naturally as the eyes gaze to the side. Aschan (1958) described studies that linked various forms of nystagmus to BAC, and Wilkinson, Kime, and Purnell (1974) reported consistent changes in Horizontal Gaze Nystagmus with increasing doses of alcohol. At the time Burns and Moskowitz were conducting their seminal research for NHTSA, Horizontal Gaze

Nystagmus recently had been found to reliably predict BACs in a study conducted in Finland (Pentilla, Tenhu, and Kataja, 1974). Further, Lehti (1976) had just calculated a strong correlation between BAC and the onset of nystagmus.

All of the field sobriety tests evaluated by Burns and Moskowitz were found to be sensitive to BAC in varying degrees, at least under laboratory conditions. In addition, all of the tests showed a consistent increase in correlations with increasing BACs. Statistical analyses found the Horizontal Gaze Nystagmus test to be the most predictive of the individual measures. However, the combined scores of two of the tests provided a slightly higher correlation than the Horizontal Gaze Nystagmus test by itself (Burns and Moskowitz, 1977); three tests were recommended to become the components of the SFST battery.

NHTSA immediately sponsored a subsequent study to standardize the test administration and scoring procedures and conduct further laboratory and field evaluations of the new battery of three tests. The researchers found that law enforcement officers tended to increase their arrest rates and were more effective in estimating the BACs of stopped drivers after they had been trained in the administration and scoring of the Standardized Field Sobriety Test battery. The results of the study were documented in the technical report, *Development and Field Test of Psychophysical Tests for DWI Arrest* (Tharp, Burns, and Moskowitz, 1981). That report was cited throughout the U.S. to establish the scientific validity of the SFST battery and to support officers' testimony in court.

Beginning in 1981, law enforcement officers used NHTSA's Standardized Field Sobriety Test (SFST) battery at roadside to help determine whether motorists who are suspected of DWI have blood alcohol concentrations (BACs) greater than 0.10 percent. Since 1981, however, many states have implemented laws that define DWI at BACs *below* 0.10. For this reason, NHTSA sponsored additional research to systematically evaluate the accuracy of the SFST battery to discriminate above or below 0.08 percent and above or below 0.04 percent BAC. In that study, Jack Stuster and Marcelline Burns (Stuster and Burns,1998) found the SFSTs to be extremely accurate. Decision analyses revealed that officers' estimates of whether a motorist's BAC was above or below 0.08 were accurate in 91 percent of the cases, and estimates of whether a motorist's BAC was above 0.04 but under 0.08 were accurate in 94 percent of the decisions to arrest and in 80 percent of the relevant cases, overall.[2]

The SFST battery is composed of three tests: Horizontal Gaze Nystagmus (HGN), Walk-and-Turn (WAT), and One-Leg Stand (OLS); the tests and scoring procedures are described in Appendix A. Table 1 compares the accuracy of the SFSTs during the 1981 and 1998 validation studies. In the 1998 study, HGN was again found to be the most accurate of the component tests in discriminating above and below the criterion BAC, and the results of the three SFSTs combined provided slightly greater accuracy than the HGN test alone. The most salient difference between the results of the 1981 and the 1998 validation studies is the substantial increase in the accuracy of officers' decisions, despite the lower criterion BAC in the 1998 study (0.10 percent BAC in 1981; 0.08 percent BAC in 1998). The greater accuracies of the SFST battery and component tests during the 1998 study are attributable to the differential experience of the officers who participated in the two studies. That is, the officers who participated in the original research had learned the procedures as part of the 1981 laboratory study; in contrast, the officers who participated in the 1998 study had been using the SFSTs for several years to help make arrest decisions under operational conditions. Thus, the levels of accuracy observed during the 1998 study reflect current conditions and should be considered the validated measures of SFST accuracy.

Table 1
Comparison of SFST Accuracy
During the 1981 and 1998 Validation Studies

	% Correct	% Correct

SFST(s)	Decisions 1981	Decisions 1998
SFST(s) 1981 1998 SFST Battery (the 3 tests combined)	81	91
Horizontal Gaze Nystagmus (HGN)	77	88
Walk-and-Turn (WAT)	68	79
One-Leg Stand (OLS)	65	83

Other studies have confirmed the considerable accuracy of the SFSTs to assist officers in making arrest decisions for DWI (Arend, et. al., 1999; Anderson and Burns, 1997; Burns and Anderson, 1995). Officers have found the SFSTs to be fully-acceptable for field use and they appreciate the diagnostic value of test results. Further, many prosecutors prefer officers to administer only the SFSTs to help make arrest decisions for DWI because the tests have been scientifically validated and are defensible in court.

NHTSA's SFSTs largely have replaced the unvalidated performance tests of unknown merit that once were the patrol officer's only tools in helping to make post-stop DWI arrest decisions. Regional and local preferences for other performance tests still exist, even though some of the tests have not been validated. Despite regional differences in what tests are used to assist officers in making DWI arrest decisions, NHTSA's SFSTs presently are used in all 50 states. NHTSA's SFSTs have become the standard pre-arrest procedures for evaluating DWI in most law enforcement agencies.[3]

> *The Horizontal Gaze Nystagmus (HGN) test is considered by many law enforcement officers to be the most effective technique to provide evidence of alcohol in a motorist's system. The normal variation in human physical and cognitive capabilities, and the effects of alcohol tolerance, can result in uncertainties when arrest decisions are made exclusively on the basis of physical and/or cognitive performance tests. These uncertainties have resulted in many DWI suspects being released rather than detained and transported to another location for evidentiary chemical testing. This is because some experienced drinkers can perform physical and cognitive tests acceptably, even with a BAC greater than 0.10 percent. However, experienced drinkers cannot conceal the physiological effects of alcohol from an officer who is skilled in HGN administration, because Horizontal Gaze Nystagmus is an involuntary reaction over which an individual has absolutely no control.*

The Importance of Standardization

The validity of SFST results is dependent upon practitioners following the established, standardized procedures for test administration and scoring. NHTSA's *SFST Student Manual* states that the procedures demonstrated in the training program describe how SFSTs should be administered under ideal conditions, but that ideal conditions do not always exist in the field. Variations from ideal conditions, and deviations from the standardized procedures, might affect the evidentiary weight that should be given to test results.

Courts in several states have reviewed the admissibility of field sobriety tests that assess physical coordination and have held that deviations in the administration of the tests should not result in the suppression of test results. These courts have found that field sobriety tests, including the Walk-and-Turn and the One-Leg-Stand of the SFST battery, are simple physical dexterity exercises that can be interpreted by an officer in the field, and by others in a court of law. However, courts have ruled that the admissibility of the HGN test may be treated differently due to its "scientific nature." For this reason, HGN results are vulnerable to challenge, and likely to be excluded by the court, if the test was not administered in strict compliance with established protocols.

Other states have been even less accommodating to deviations from the standardized procedures. In particular, the Ohio State Supreme Court ruled that law enforcement officers have no discretion in the administration of SFSTs. In a four-to-two decision, the Ohio State Supreme Court held in *Ohio v. Homan,* 732 N.E.2d 952 (Ohio 2000), that Standardized Field Sobriety Tests conducted in a manner that departs from the methods established by NHTSA "are inherently unreliable" and thus inadmissible.[4]

The SFST battery is composed of three separate tests with three independent predictive validities that range from 79 to 88 percent. Depending on the physical characteristics of the subject and roadside conditions, an officer might choose to refrain from administering the entire SFST battery, as directed by the training materials (e.g., a leg injury that might affect a person's ability to perform the OLS test). Because an officer is permitted the discretion to withhold a test, it is reasonable to question why a deviation in the administration of one of the three tests would disqualify the entire battery. Although it is not recommended to do so under ideal conditions, the data show that accurate arrest decisions reliably can be made on the basis of two of the SFSTs, or on the basis of HGN test results, alone.

The International Association of Chiefs of Police (IACP) adopted uniform procedures in 1992 to guide the training of SFST instructors and practitioners. Those standards include 24-hours of NHTSA-approved SFST instruction. The procedures for administering and interpreting SFST results can be readily learned and, generally, proficiency increases with experience. However, it is possible for SFST skills to degrade if they are not exercised regularly (e.g., during a prolonged absence from patrol work). Also, the SFST procedures have evolved since they were first developed in 1981. Modifications to the standardized procedures could result in an officer administering SFSTs according to outdated protocols.[5] For these reasons, NHTSA recommends that law enforcement agencies conduct refresher training for SFST instructors and practitioners.

Project Objectives

The primary objective of this study is to develop a model system to help law enforcement agencies manage Standardized Field Sobriety Test (SFST) training requirements. A further objective is to explore the feasibility of establishing and operating a statewide SFST training records system.

General Approach

Judges in the State of Colorado became concerned with inconsistencies in the testimony of law enforcement officers concerning SFST administration and scoring procedures. In response to those concerns, representatives of law enforcement agencies, the Rocky Mountain Institute for Transportation Safety, and the Colorado Department of Transportation developed standards for SFST instructors and practitioners, based on the NHTSA standards, which include requirements for refresher training. In this regard, the Colorado SFST standards require that practitioners receive at least two hours of refresher training every two years and instructors receive at least eight hours of refresher training every two years, to maintain their SFST practitioner and instructor certifications. The statewide regulation took effect in 1999, with a two-year grandfather clause expiring in November of 2001.

The implementation of SFST refresher training requirements by the State of Colorado offers an

opportunity to study how law enforcement agencies maintain records of training experience to comply with the requirement. The question of particular interest is, how do agencies identify when individual SFST instructors and practitioners must receive their periodic refresher training? Interviews were conducted with personnel from a sample of Colorado law enforcement agencies to obtain the information necessary to answer the research questions.

---

1. Various terms are used throughout the United States for offenses involving drinking and driving. In this report, Driving While Impaired (DWI) is used to refer to all occurrences of driving at or above the illegal blood alcohol concentratiion (BAC) limit of a jurisdiction.
2. In addition to the results of the decision analysis, the study found statistically significant correlations between SFST results and measured BACs (p=.005); also, the difference between the mean estimated and measured BACs of the 297 motorists tested at roadside during the field study was very small and operationally irrelevant (i.e., 0.117 vs. 0.122 percent BAC, respectively).
3. The Advisory Committee on Highway Safety of the International Association of Chiefs of Police (IACP) recommended in 1986 that law enforcement agencies adopt and implement NHTSA's SFSTs and the associated training program.
4. Officers always should fully comply with NHTSA's guidelines when administering the SFSTs. However, if deviations occur, officers and the courts should understand that any deviation from established procedures relates to the *weight* of the evidence, not its *admissibility*.
5. For example, the original SFST procedures specified that the HGN test not be administered to individuals who were wearing hard contact lenses. The stipulation was made to avoid the possibility of losing a lens as a consequence of the required eye movements. The stipulation eventually was removed when it was recognized that the possibility of dislodging a contact lens was minimal.

- **Continue to Research**
- **Table of Contents**